

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: August 18, 2002, 08:18:46 : Search time 61.25 Seconds  
(without alignments)  
865.016 Million cell updates/sec

Title: US-09-777-921A-2  
Perfect score: 2491  
Sequence: 1 MLRWLRDFALPTAACQDAEQ.....VGISVYVENMKOTLGVTKQ 477

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2471	99.6	477	22	AAW79077 Human protein SEQ
2	1552.5	62.6	471	22	AAU27697 Human full-length
3	1552.5	62.6	509	22	AAU27869 Human contig poly
4	1501	60.5	469	22	AAU66718 Membrane-bound pro
5	1501	60.5	469	22	AAU67554 Human PRO1106. HO
6	1501	60.5	469	22	AAU65241 Human PRO1106 (UNQ
7	1499	60.4	469	22	AAU76084 Murine ADP/ATP tra
8	1499	60.4	469	22	AAU56023 Skin cell protein,
9	1479.5	59.6	508	22	ABG22637 Novel human diapo
10	1394	56.2	385	21	AAU42129 Human ORFX ORF2093
11	1325	53.4	366	22	AAU40072 Human polypeptide

12	1241	50.0	461	22	AAU5800 Human membrane tra
13	1143	46.1	244	22	AAU43524 Human polypeptide
14	1143	46.1	244	22	AAU19941 Novel human calciu
15	1037	41.8	292	21	AAU50388 Human uncoupling p
16	910	36.7	312	21	AAU75964 Murine skin cell p
17	910	36.7	312	22	AAU55903 Skin cell protein,
18	902	36.4	370	22	ABU71306 Drosophila melanog
19	841	33.9	182	22	AAU80051 Human protein SEQ
20	800.5	32.3	226	22	AAU41858 Human polypeptide
21	761	30.7	208	22	AAU93247 Human polypeptide
22	632	25.5	169	21	AAU50389 Human uncoupling p
23	521	21.0	352	21	AAU30070 Arabidopsis thalia
24	508	20.5	241	21	AAU12158 Arabidopsis thalia
25	508	20.5	330	21	AAU28434 Arabidopsis thalia
26	508	20.5	332	21	AAU28433 Arabidopsis thalia
27	505.5	20.4	166	21	AAU42319 Human ORFX ORF2083
28	479	19.3	138	22	ABG22634 Novel human diapo
29	455.5	18.4	415	21	AAU29005 Arabidopsis thalia
30	454.5	18.3	289	20	AAU31936 Wheat brittle-1 pa
31	454.5	18.3	381	21	AAU29257 Arabidopsis thalia
32	454.5	18.3	384	21	AAU29256 Arabidopsis thalia
33	446	18.0	392	21	AAU11515 Arabidopsis thalia
34	446	18.0	392	21	AAU49411 Arabidopsis thalia
35	446	18.0	411	21	AAU11514 Arabidopsis thalia
36	446	18.0	411	21	AAU49410 Arabidopsis thalia
37	445	17.9	291	21	AAU30071 Arabidopsis thalia
38	432.5	17.4	316	21	AAU11516 Arabidopsis thalia
39	432.5	17.4	316	21	AAU49412 Arabidopsis thalia
40	417	16.8	198	21	AAU12159 Arabidopsis thalia
41	412	16.6	266	21	AAU28435 Arabidopsis thalia
42	410.5	16.5	377	22	ABU60506 Drosophila melanog
43	382.5	15.4	273	21	AAU29006 Arabidopsis thalia
44	380.5	15.3	267	21	AAU29258 Arabidopsis thalia
45	369	14.9	323	22	AAU41142 Human polypeptide

## ALIGNMENTS

## RESULT 1

AAW79077  
ID AAW79077 standard; Protein: 477 AA.

XX AC AAW79077;

XX DT 06-NOV-2001 (first entry)

XX DE Human protein SEQ ID NO 1739.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 30-OCT-2000; 2000US-0693325.

XX PA (HYSE-) HYSEQ INC.



Matches 286; Conservative 85; Mismatches 87; Indels 7; Gaps 2;

Oy 17 DAEOPTRYETLFOALDRNGGVVDIGELQBLRNLGIPILGQD----AEEKIFTTGDVNRK 72  
 Db 7 daerqrwgrlfealdankdgrvvhelqglarlg---ggnpdpgagqgissegdadp 63  
 Oy 73 KGLDFEFMYLKDHEKMKLAFKSLDKNDGKIEASEIVQSLOTGLTISQOAEILQ 132  
 Db 64 ggldeafsyrlqereqrlllmfhsldrnqghidvseiqsfalgisiseqaelkh 123  
 Oy 133 SIDVGTWTVDWNEWRDYFTFNPVTIEEIRFWKHSTGIDIGSLTIPDEFTDEKSG 192  
 Db 124 smdrdgtmdidwqewdhflhslenvedvlyfwkshstvdldgecltvpdefskqektg 183  
 Oy 193 QMWRQLLAGGIAGAVSRTSTAPDLRLKIMQVHGSKSDKMNIFGFRWVHGGRSLWR 252  
 Db 184 mwwqlvagavagavartgtcapldrlkvfmqvhasktnrlnlgslrsmviegslr 243  
 Oy 253 NGTNTVTKIAPETAVKFWAYEQYKLLTBEGQKIGTFERFISGSMAGATAQTFTYPMV 312  
 Db 244 gnginviklapesaifmayeqikrallgqgetlhvqerfvagslagataqtilypmevl 303  
 Oy 313 KTRLAVGKTQYSGIYDCAKILKHEGLGAFYKGVVNPVNLGIPVAGIDLAVYELLKSYW 372  
 Db 304 ktrltlrtrtqykgllidcarrileregprafyrgyilpvnvlgilpvgidlavetlknw 363  
 Oy 373 LDNFAKDSVNPVWVLLGCGALSSTCGQLASYPALVTRRMOAQLGSGPOLNNVGLFR 432  
 Db 364 lqyghdsadpgilvilacgtisastcgqlasyplalvtrmqagaleggpqlsmlgllr 423  
 Oy 433 RIISKEGIPGLRGITPNNPKVLPVAGISYVYVYNNKOTLGVTK 477  
 Db 424 hllsqgmrglyrglapnmkvpavlsyvyvymkqelgytar 468

RESULT 3

AAU27869  
 ID AAU27869 standard; Protein: 509 AA.

XX AAU27869;

XX 18-DEC-2001 (first entry)

XX Human contig polypeptide sequence #22.

XX Mammal: human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;  
 KW nervous system disorder; inflammatory disorder; cell differentiation;  
 KW angiogenesis; stem cell growth factor; actinin; inhibin; cartilage; burn;  
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
 KW cytoskeletal; antitumor; antirheumatic; antiarthritic; vulnary; antinflammatory;  
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;  
 KW neuroprotective; osteopathic; antidiabetic; antiallergic;  
 KW immunostimulant; analgesic; gene therapy.

XX Homo sapiens.

OS Synthetic.

XX WO200164834-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04936.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PR 17-JUN-2000; 2000US-0507707.

PR 14-JUL-2000; 2000US-0616807.

PR 19-SEP-2000; 2000US-0664641.

XX (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Mehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
 PI Drmanac R;  
 DR WPI: 2001-589862/66.  
 DR N-PSDB: AAS44769.  
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries  
 PT prepared from various human tissues, for diagnosis, treatment of  
 PT cancer, neurological, inflammatory disorders and for use in arrays for  
 PT detection  
 XX  
 PS Claim 10: Page 126-127: 153pp: English.

XX Sequences AAU27676-AAU28019 represent full-length polypeptides and  
 CC contig polypeptides of the invention. The proteins and their associated  
 CC DNA sequences are useful for the treatment, diagnosis and prevention of  
 CC various types of disorder in a mammalian subject such as a human, dog,  
 CC monkey, mouse, hamster or rat. The disorders include cancers such as  
 CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as  
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,  
 CC cell proliferation, cell differentiation, stem cell growth factor,  
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells  
 CC in culture to give rise to neuroepithelial cells that can be used to  
 CC augment or replace cells damaged by illness, accidental damage or genetic  
 CC disorders. The sequences may also be used for regeneration of bone,  
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.  
 CC Note: Some sequences for this patent did not form part of the printed  
 CC specification, but were obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 509 AA;

Query Match 62.6%; Score 1552.5; DB 22; Length 509;  
 Best Local Similarity 61.5%; Pred. No. 2.6e-133;  
 Matches 286; Conservative 85; Mismatches 87; Indels 7; Gaps 2;

Oy 17 DAEOPTRYETLFOALDRNGGVVDIGELQBLRNLGIPILGQD----AEEKIFTTGDVNRK 72  
 Db 45 daerqrwgrlfealdankdgrvvhelqglarlg---ggnpdpgagqgissegdadp 101  
 Oy 73 KGLDFEFMYLKDHEKMKLAFKSLDKNDGKIEASEIVQSLOTGLTISQOAEILQ 132  
 Db 102 ggldeafsyrlqereqrlllmfhsldrnqghidvseiqsfalgisiseqaelkh 161  
 Oy 133 SIDVGTWTVDWNEWRDYFTFNPVTIEEIRFWKHSTGIDIGSLTIPDEFTDEKSG 192  
 Db 162 smdrdgtmdidwqewdhflhslenvedvlyfwkshstvdldgecltvpdefskqektg 221  
 Oy 193 QMWRQLLAGGIAGAVSRTSTAPDLRLKIMQVHGSKSDKMNIFGFRWVHGGRSLWR 252  
 Db 222 mwwqlvagavagavartgtcapldrlkvfmqvhasktnrlnlgslrsmviegslr 281  
 Oy 253 NGTNTVTKIAPETAVKFWAYEQYKLLTBEGQKIGTFERFISGSMAGATAQTFTYPMV 312  
 Db 282 gnginviklapesaifmayeqikrallgqgetlhvqerfvagslagataqtilypmevl 341  
 Oy 313 KTRLAVGKTQYSGIYDCAKILKHEGLGAFYKGVVNPVNLGIPVAGIDLAVYELLKSYW 372  
 Db 342 ktrltlrtrtqykgllidcarrileregprafyrgyilpvnvlgilpvgidlavetlknw 401  
 Oy 373 LDNFAKDSVNPVWVLLGCGALSSTCGQLASYPALVTRRMOAQLGSGPOLNNVGLFR 432  
 Db 402 lqyghdsadpgilvilacgtisastcgqlasyplalvtrmqagaleggpqlsmlgllr 461  
 Oy 433 RIISKEGIPGLRGITPNNPKVLPVAGISYVYVYNNKOTLGVTK 477

Db 462 hilsqegmrglyrglapnfmkvpavslayvvvymkqalgvtar 506

RESULT 4

AAV66718  
ID AAY66718 standard; protein: 469 AA.

XX AC AAY66718;

XX DT 05-APR-2000 (first entry)

XX DE Membrane-bound protein PRO1106.

XX KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;

XX KW Pharmaceutical; receptor immunoadhesin; gene mapping.

XX OS Homo sapiens.

XX PN WO9963088-A2.

XX PD 09-DEC-1999.

XX PF 02-JUN-1999; 99WO-US12252.

XX PR 02-JUN-1998; 98US-0087607.

XX PR 02-JUN-1998; 98US-0087609.

XX PR 03-JUN-1998; 98US-0087759.

XX PR 04-JUN-1998; 98US-0087827.

XX PR 04-JUN-1998; 98US-0088021.

XX PR 04-JUN-1998; 98US-0088025.

XX PR 04-JUN-1998; 98US-0088028.

XX PR 04-JUN-1998; 98US-0088029.

XX PR 04-JUN-1998; 98US-0088030.

XX PR 04-JUN-1998; 98US-0088033.

XX PR 05-JUN-1998; 98US-0088326.

XX PR 05-JUN-1998; 98US-0088167.

XX PR 05-JUN-1998; 98US-0088202.

XX PR 05-JUN-1998; 98US-0088212.

XX PR 05-JUN-1998; 98US-0088217.

XX PR 10-JUN-1998; 98US-0088655.

XX PR 10-JUN-1998; 98US-0088722.

XX PR 10-JUN-1998; 98US-0088730.

XX PR 10-JUN-1998; 98US-0088734.

XX PR 10-JUN-1998; 98US-0088738.

XX PR 10-JUN-1998; 98US-0088740.

XX PR 10-JUN-1998; 98US-0088741.

XX PR 10-JUN-1998; 98US-0088742.

XX PR 10-JUN-1998; 98US-0088810.

XX PR 10-JUN-1998; 98US-0088811.

XX PR 10-JUN-1998; 98US-0088824.

XX PR 10-JUN-1998; 98US-0088825.

XX PR 11-JUN-1998; 98US-0088826.

XX PR 11-JUN-1998; 98US-0088858.

XX PR 11-JUN-1998; 98US-0088861.

XX PR 11-JUN-1998; 98US-0088863.

PR 19-JUN-1998; 98US-0089952.  
PR 22-JUN-1998; 98US-0090246.  
PR 22-JUN-1998; 98US-0090252.  
PR 22-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 23-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 24-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090678.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090690.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091626.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 07-JUL-1998; 98US-0091982.  
PR 09-JUL-1998; 98US-0092182.  
PR 10-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.  
PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095318.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095916.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.  
PR 12-AUG-1998; 98US-0096329.  
PR 17-AUG-1998; 98US-0096757.  
PR 17-AUG-1998; 98US-0096766.  
PR 17-AUG-1998; 98US-0096768.  
PR 17-AUG-1998; 98US-0096773.  
PR 17-AUG-1998; 98US-0096791.  
PR 17-AUG-1998; 98US-0096867.  
PR 17-AUG-1998; 98US-0096891.  
PR 17-AUG-1998; 98US-0096894.  
PR 17-AUG-1998; 98US-0096895.  
PR 17-AUG-1998; 98US-0096897.  
PR 18-AUG-1998; 98US-0096949.  
PR 18-AUG-1998; 98US-0096950.  
PR 18-AUG-1998; 98US-0096959.  
PR 18-AUG-1998; 98US-0096960.  
PR 18-AUG-1998; 98US-0097022.  
PR 19-AUG-1998; 98US-0097141.

PR 20-AUG-1998: 98US-0097218.  
 PR 24-AUG-1998: 98US-0097661.  
 PR 26-AUG-1998: 98US-0097951.  
 PR 26-AUG-1998: 98US-0097952.  
 PR 26-AUG-1998: 98US-0097953.  
 PR 26-AUG-1998: 98US-0097955.  
 PR 26-AUG-1998: 98US-0097971.  
 PR 26-AUG-1998: 98US-0097974.  
 PR 26-AUG-1998: 98US-0097978.  
 PR 26-AUG-1998: 98US-0097979.  
 PR 26-AUG-1998: 98US-0097986.  
 PR 26-AUG-1998: 98US-0098014.  
 PR 31-AUG-1998: 98US-0098525.  
 PR 16-SEP-1998: 98US-0100634.  
 PR 12-JAN-1999: 99US-0115565.  
 XX (GETH ) GENE/TECH INC.  
 PA Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
 PI Wood WI, Yuan J;  
 PI WPI: 2000-072883/06.  
 DR N-PSDB; AA265058.  
 XX Membrane-bound proteins and related nucleotide sequences -  
 XX claim 12; Fig 206; 822pp; English.  
 XX The invention provides membrane-bound PRO polypeptides and  
 CC polynucleotides encoding them. The PRO sequences of the invention were  
 CC identified based on extracellular domain homology screening. The PRO  
 CC sequences have homology with proteins including LDL receptor, TIE  
 CC ligands and various enzymes. The membrane-bound proteins and receptor  
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
 CC immunoadhesins, for instance, can be used as therapeutic agents to block  
 CC receptor-ligand interactions. The membrane-bound proteins can also be  
 CC employed for screening of potential peptide or small molecule inhibitors  
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
 CC are useful as hybridization probes, in chromosome and gene mapping and in  
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
 CC will also be useful for the preparation of PRO polypeptides, especially  
 CC by recombinant techniques.  
 XX Sequence 469 AA;

Query Match 60.5%; Score 1501; DB 21; Length 469;  
 Best Local Similarity 65.8%; Pred. No. 1.2e-128;  
 Matches 275; Conservative 71; Mismatches 72; Indels 0; Gaps 0;  
 QY 60 EEKIFTGDNVKGKLDPEEFMKYLDHEKMKLAFKSLDKNDGKTEASEIVQSLOTGLG 119  
 DB 52 kqkivagdkldgldfeefvhyldhekkrlvfkldkndgridagqmsldrlg 111  
 QY 120 LRISEQAEELIQLSDVDGTVTDNENWROFLENPNVTDIEIRFKWKNSTGIDGSLT 179  
 DB 112 vkiseqaeelkikmdkngntidvnewrdyhlhlpvenipellykwhstfidvgenit 171  
 QY 180 IPDEFTEDEKSGQWROLLAGGAGAVSRTSTAPDLKIMQVHGSKDKMNIFGGFR 239  
 DB 172 vpdftveerqtgmwrthlvagggagavstctcpldrklvlnqvhaarsnmglvggft 231  
 QY 240 QMVKEGGIRSLWRNGTAVKVIAPETAVKFWAYEYOYKLLTEGQKIGTFERFISGMAG 299  
 DB 232 qmireggarslwrnglnvliapetavkfwayeqkrlvgsdqetirherlvagslag 291  
 QY 300 ATAQTFYPMEMVMTRLAVGVTGQYSGIYDCAKKILKHEGLGAFYKGYVNPNLGIIPVAG 359  
 DB 292 alagsiypmevltrmalrktgysgmldcarrilaregvaafykygypnmglipvag 351  
 QY 360 IDLAYETELKSGYLDNDKSNVPMVLLCGGALSSTCGLASYPALVTRMQAQL 419  
 DB 352 idlayetelkswlqhvyavnsadpgvfvllacgtmstcqqiasypialvtrmqaqasi 411

OY 420 EGSPOLNMVGLFERRIISKEGIPYRGITPNFMKVLPAVGISYVYVENKMTGLGVTK 477  
 DB 412 egapevtmsslfkhlrtgafgiyrglapnfmkvlpavslayvvyenlkltlgvqsr 469

RESULT 5  
 ID AAB87554 standard; Protein; 469 AA.  
 AC AAB87554;  
 XX 15-MAY-2001 (first entry)  
 DT Human PRO1106.  
 DE Human; PRO protein; mapping.  
 XX Homo sapiens.  
 XX WO200116318-A2.  
 XX 08-MAR-2001.  
 XX 24-AUG-2000; 2000WO-US23328.  
 PR 01-SEP-1999; 99WO-US20111.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 07-DEC-1999; 99US-0169495.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 11-JAN-2000; 2000US-0175481.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 03-MAR-2000; 2000US-0187202.  
 PR 25-APR-2000; 2000US-0199397.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 05-JUN-2000; 2000US-0209832.  
 XX (GETH ) GENE/TECH INC.  
 PA Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;  
 PI WPI: 2001-183260/18.  
 DR N-PSDB; AAF32086.  
 XX Eighty four nucleic acids encoding PRO polypeptides, useful in  
 PT molecular biology, including use as hybridization probes, and in  
 XX chromosome and gene mapping.  
 XX Claim 12; Fig 58; 278pp; English.

The present sequence is a human PRO polypeptide (secreted and transmembrane). The PRO protein, and PRO agonists, PRO antagonists or anti-PRO antibodies are useful for preparation of a medicament useful in the treatment of a condition which is responsive to the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO protein may also be employed as molecular weight markers for protein electrophoresis. The PRO coding sequence has applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping.

Sequence 469 AA;

Query Match 60.5%; Score 1501; DB 22; Length 469;  
 Best Local Similarity 65.8%; Pred. No. 1.2e-128;  
 Matches 275; Conservative 71; Mismatches 72; Indels 0; Gaps 0;

OY 60 EEKIFTGDNVKGKLDPEEFMKYLDHEKMKLAFKSLDKNDGKTEASEIVQSLOTGLG 119  
 DB 52 kqkivagdkldgldfeefvhyldhekkrlvfkldkndgridagqmsldrlg 111

QY 120 LTISEQOAEILQSIDVDGTMVDNENRDFLEPNVDTIEEIRFWKHSTGIDIGDSL 179  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 112 vkiseqaeiklsmdkngtmdidnewrdyhlhnpvenipeilfykxhstfidvgenlt 171  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 180 IPDETEDEKSGOWHROLLAGGTAGAVSRSTAPDLRLKIMQVHGSKSDKNISGGR 239  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 172 vpedtveerqcmwrlhvgggagavsrctctaplrlkvlmgvhasrnmngivggft 231  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 240 QMVKEGIRSLWRGNGTNYIKIAPETAVKFWAYEQYKLLTEBQKIGTGFERSGSMAG 299  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 232 qmireggarslwrnginvlkapesaalkmayeqikrlvsgdgetlrlherlvagslag 291  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 300 ATAQTFTYPMVKNTRLAVGKTGOYSGIYDCAKKILKHGEGAFYKGYVNPNLGIIPYAG 359  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 292 alaqeslypmevlktmrlkrtgqysgmldccarrilaregvaafykyvnpnmgliipyag 351  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 360 IDLAVYELLSYWLDFNFAKDSVNPVGVVLLCGGALSSTCGQLASYPALVTRTMOAQL 419  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 352 idlavyetlknawlhqyavnsdpvgfvllacgtmsstcgqlasypalvtrtmaqasi 411  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 420 EGSPQLNMVGLFRRIISKEGIPGLYRGITPNFMKVLPAVGISYVYENNKOTLGVTK 477  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 412 egapvmtsslfkhlrttegafglyrglapnfmkvpavsisvvyenkitigvqsr 469  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|

RESULT 6  
 AAB65241 ID AAB65241 standard; Protein: 469 AA.  
 XX AC AAB65241;  
 XX DT 02-APR-2001 (first entry)  
 XX DE Human PRO1106 (UNQ549) protein sequence SEQ ID NO:289.  
 XX KW Human: secreted and transmembrane protein; PRO: cytosolic;  
 KW cell death; cancer; chromosomal mapping: gene mapping; tissue typing;  
 KW diagnostic assay.  
 XX OS Homo sapiens.  
 XX PN WO200073454-A1.  
 XX PD 07-DEC-2000.  
 XX PF 30-MAR-2000; 2000WO-US08439.  
 XX PR 02-JUN-1999; 99WO-US12252.  
 PR 23-JUN-1999; 99US-0141037.  
 PR 20-JUL-1999; 99US-0143048.  
 PR 26-JUL-1999; 99US-0144758.  
 PR 28-JUL-1999; 99US-0145698.  
 PR 17-AUG-1999; 99US-0149396.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 PR 08-OCT-1999; 99US-0158663.  
 PR 30-NOV-1999; 99WO-US28313.  
 PR 16-DEC-1999; 99WO-US28301.  
 PR 20-DEC-1999; 99WO-US30095.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 02-MAR-2000; 2000WO-US05004.  
 PR 15-MAR-2000; 2000WO-US05841.  
 PR 20-MAR-2000; 2000WO-US06884.  
 XX 20-MAR-2000; 2000WO-US07377.

(GETH ) GENENTECH INC.  
 Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
 Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
 Grimaldi CJ, Gurney AL, Kijavini IJ, Napier MA, Pan J, Paoni NF;  
 Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
 Zhang Z;  
 WPI: 2001-032160/04.  
 DR N-PSDB: AAF44204.  
 XX PRO polynucleotides used to produce polypeptides used to target  
 PT bioactive molecules such as toxins, radiolabels or antibodies, to  
 PT specific cells, to cause targeted cell death -  
 XX Claim 12: Fig 206; 935pp; English.  
 XX The present invention describes human secreted and transmembrane PRO  
 CC proteins. The PRO proteins have cytostatic activity. The PRO proteins  
 CC can be used for targeted delivery of bioactive molecules, such as  
 CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
 CC sequences, and their fragments, can be used as hybridisation probes, in  
 CC chromosomal and gene mapping, and in the generation of anti-sense RNA  
 CC and DNA. They may also be used to produce transgenic animals which are  
 CC used to develop and screen therapeutically useful reagents. The PRO  
 CC nucleotide and protein sequence can be used for tissue typing and in  
 CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.  
 CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used  
 CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and  
 CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein  
 CC sequences given in the exemplification of the present invention.  
 XX SQ Sequence 469 AA;  
 Query Match 60.5%; Score 1501; DB 22; Length 469;  
 Best Local Similarity 65.8%; Pred. No. 1.2e-128;  
 Matches 275; Conservative 71; Mismatches 72; Indels 0; Gaps 0;  
 QY 60 EEKIFTTGDVNRDGLDPEEFNKYLKDHKKMLAFKSLDKNDKIEASEIVQSQTGL 119  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 52 kqkivqgdgldgldfeefvhyldhekkrlrvfkldkndgridaqemqslrdlg 111  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 120 LTISEQOAEILQSIDVDGTMVDNENRDFLEPNVDTIEEIRFWKHSTGIDIGDSL 179  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 112 vkiseqaeiklsmdkngtmdidnewrdyhlhnpvenipeilfykxhstfidvgenlt 171  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 180 IPDETEDEKSGOWHROLLAGGTAGAVSRSTAPDLRLKIMQVHGSKSDKNISGGR 239  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 172 vpedtveerqcmwrlhvgggagavsrctctaplrlkvlmgvhasrnmngivggft 231  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 240 QMVKEGIRSLWRGNGTNYIKIAPETAVKFWAYEQYKLLTEBQKIGTGFERSGSMAG 299  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 232 qmireggarslwrnginvlkapesaalkmayeqikrlvsgdgetlrlherlvagslag 291  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 300 ATAQTFTYPMVKNTRLAVGKTGOYSGIYDCAKKILKHGEGAFYKGYVNPNLGIIPYAG 359  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 292 alaqeslypmevlktmrlkrtgqysgmldccarrilaregvaafykyvnpnmgliipyag 351  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 360 IDLAVYELLSYWLDFNFAKDSVNPVGVVLLCGGALSSTCGQLASYPALVTRTMOAQL 419  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 352 idlavyetlknawlhqyavnsdpvgfvllacgtmsstcgqlasypalvtrtmaqasi 411  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 QY 420 EGSPQLNMVGLFRRIISKEGIPGLYRGITPNFMKVLPAVGISYVYENNKOTLGVTK 477  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|  
 Db 412 egapvmtsslfkhlrttegafglyrglapnfmkvpavsisvvyenkitigvqsr 469  
 : ||||| ||:| :|||:||||| : ||: ||| :||| :||:|

RESULT 7  
 AAY76084 ID AAY76084 standard; Protein: 469 AA.  
 XX XX  
 AC AAY76084;





Db 52 kqkivagdkldgldfeefvhyldhekkrlrvfksldkndgridaeqmqlrdlg 111  
 QY 120 LTISEOAEILIQSDIVGTWTDVDMNWRDYFLNPNVTDIEEIRFKWSTGTGIDGSTW 179  
 Db 112 vxiseqqeeklikamdngatidwnewrdyhlhprvenpeillykwhstfidvgenit 171  
 QY 180 IDEPTEDEKSKQWROLLAGGTAGAVSRTSTAPLRLKIMQVHGSKSDKNIFGGR 239  
 Db 172 vpedtveerqgmwrhnlvaggagavsrctcpldrklvlnqvhassnmcivgft 231  
 QY 240 QMVKEGIRSLWRGNTVINIAPETAVKFWAYEYOYKLLTEGQKIGTFPERFISGMAG 299  
 Db 232 qmireggakslvrgnglnvklapesaalkfmayeqmrlvgsdgetlrherlivagslag 291  
 QY 300 ATAOTFIYPMEVMTKRLAVGKTGOYSGIYDCAKKILKHEGLGAFYKYVNPULLGIIPYAG 359  
 Db 292 alaqsilypmeviktrmairktgysgmldcarrilakegvaafykygipnmlgiipyag 351  
 QY 360 IDLAVYELLSKWLONFADSVNPGVMVLLGCGALSTCCQLASYPLALVTRMQAAML 419  
 Db 352 idlavyetlknwtlgrvavnsadpgvfvllacgtlscqqlasyplalvtrmqaqasi 411  
 QY 420 EGSPLNVMVGLFRRIISKEGIPGLYRGITPNEMKVLPAVGISYVYENMKOTLGVYOK 477  
 Db 412 egapevtnssifkqlrlrtcgafglyrglapnfmkvpavsisyvyenmkilvgvqr 469

## RESULT 9

ABG22637  
 ID ABG22637 standard; Protein: 508 AA.

AC ABG22637

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #22628.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175057-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB: RAS86824.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity

PS Claim 20; SEQ ID No 52996; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 and gene mapping, and in recombinant production of (II). The  
 polynucleotides are also used in diagnostics as expressed sequence tags  
 for identifying expressed genes. (I) is useful in gene therapy techniques  
 to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or  
 quantitating a polypeptide in tissue, as molecular weight markers and as  
 a food supplement. (II) and its binding partners are useful in medical  
 imaging of sites expressing (II). (I) and (II) are useful for treating  
 disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 508 AA;

Query Match 59.6%; Score 1479.5; DB 22; Length 508;  
 Best Local Similarity 59.3%; Pred. No. 1.3e-126;  
 Matches 277; Conservative 85; Mismatches 96; Indels 9; Gaps 4;

QY 17 DAOPRTYETLQALDRNGDGVWDIGELQELRNGLPLGQD----AEKIFTTGDVNDK 72

Db 45 dhertrgrvrlfeeldsnkdgrvvhelrqglarlg---ggnpdpgaggissegdapn 101

QY 73 GKLDFFEFMKYLDKHEKKMLAFKSLDRKNDGKTEASEIVQSLQTLGLTISEQOAE-IL 131

Db 102 ggldleefarylqereqrlllmfaldnrgqghldvseiqqsfraigisilleqaeknfa 161

QY 132 QSTDVDGTMVDNMWRDYELFNPVTDIEEIRFKWSTGTGIDGSTLTPDETEDEKRS 191

Db 162 qavrdgtctidwgdwrhflhseinvvedviyfwkhnstvidgecltvpdelskgekit 221

QY 192 GQWNROLLAGGIAGAVSRTSTAPLRLKIMQVHGSKSDKNIFGGRQWKEGGIRSLW 251

Db 222 gmwwkqlaagavagavavsrctcpldrklvfmqvhasknrlnllgslxsmvleggirsrl 281

QY 252 RGNQNVNVIKIPETAVKFWAYEYOYKLLTEGQKIGTFPERFISGMAGATAOTFIYPMEV 311

Db 282 rgnginvklapesaalkfmayeqikraipqeketlpvlerfvagslagataqtilypmav 341

QY 312 MKTRLAVGKTGOYSGIYDCAK-KILKHEGLGAFYKYVNPULLGIIPYAGIDLAVYELLS 370

Db 342 lkrtrtirttcgqykglaglrrg9lleregprafyrgyipnvgilipyagidlavyetlkn 401

QY 371 YWLDNFAKDSVNPVGMVLLGCGALSTCCQLASYPLALVTRMQAAMLEGSPQLMAYGL 430

Db 402 wwlgqyshdsadpgllvllacgtlscqqlasyplalvtrmqaqasleggpklsamlgl 461

QY 431 FRRIISKEGIPGLYRGITPNEMKVLPAVGISYVYENMKOTLGVYOK 477

Db 462 lrhlilsqegmrglyrglapnfmkvpavsisyvyenmkqalgvtar 508

## RESULT 10

AAAB42329

ID AAAB42329 standard; Protein: 385 AA.

AC AAAB42329;

DT 08-FEB-2001 (first entry)

DE Human ORF2093 polypeptide sequence SEQ ID NO:4186.

KW Human; open reading frame; ORF; detection; cytostatic; hepatotropic;  
 KW vulnery; antipsoaric; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;



KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.

OS Homo sapiens.

XX W0200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000MO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI: 2000-602362/57.

XX N-PSDB; AAC76538.

XX Novel nucleic acids and peptides derived from open reading frame X,  
 useful for treating e.g. cancers, proliferative disorders,  
 neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 3378-3379; 5507pp; English.

XX AAC7446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,  
 which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 antiproliferative; antiparkinsonian; nontropic; neuroprotective;  
 osteopathic; anticonvulsant; antihypertensive; immunosuppressant;  
 immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;  
 antidiabetic; hypotensive; dermatological; antifungal; antirheumatic;  
 antiinflammatory; antibacterial; antiviral; antifungal; antihypertensive;  
 antithyroid; and antianemic. The sequences can be used for determining  
 the presence of or predisposition to, or preventing or treating  
 pathological conditions associated with an ORFX-associated disorder. The  
 nucleic acids can be used to express ORFX proteins in gene therapy  
 vectors. The proteins and nucleic acids may be used to treat cancers,  
 proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 graft vs host disease, cardiovascular disease, diabetes mellitus,  
 hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 385 AA;

Query Match 56.28; Score 1394; DB 21; Length 385;  
 Best Local Similarity 66.84; Pred. No. 5.6e-119;  
 Matches 256; Conservative 63; Mismatches 64; Indels 0; Gaps 0;

QY 95 FKSLLKNDGKEAEIVQSLOTGLTISQQAELILQSDVDTMTVDNENRDFLEN 154

DB 3 fksldkndgridaqlmqslrdlgvklseqgaclksmdkngtmdcldnwdrynlh 62

QY 155 PVTDEEIRFWKSTGTIDGDSLTIPDETEDEKKSQGWHROLLAGTAGVSRSTAP 214

DB 63 pvenipellywkhstfdvgenltvpdefteerqgmwrhfvagggagavsrctap 122

QY 215 LDRKIMQVHGSKDKNIFGSGRQWVKEGGIRSLRNGTNNIKIAPETAVKFWAYE 274

DB 123 ldrklvlnqvhsrtnnmglvggtqmqreggarslrvrgnglnvkiapesaikfmayeq 182

QY 275 YKLLTEBQCKIGTERTFISGSMAGATAOTIYPAEWNKRLAVGTGQYSGIYCAKKI 334  
 DB 183 lkrlyvsdqetlrirhivagslagalaqsslypmeviktrmalrktgysgmldcarrl 242  
 QY 335 LKHEGLGAFYKGYVPENLGIIPYAGIDLAVALYELKSYWLNFAKDSVNPQVYLLCGGAL 394  
 DB 243 laregvaafykygypnmglipagidlavyetlknawlhqyavnsedpavfllacgm 302  
 QY 395 SSTCGOLASYPLALVBTMOAQLGSPOLNWNGLFRRISKEGIPGLYRGTPPMKV 454  
 DB 303 stscqslasyplalvtrtmqasiegepvcmslfxhrlrtegafgyrglagnfmkv 362  
 QY 455 LPVAGISYVYENKQTLGVTK 477  
 DB 363 lpavaisyvyenykitlgvqr 385

RESULT 11

AAM40072

XX AAM40072 standard; Protein; 366 AA.

XX AAM40072;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3217.

XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX Homo sapiens.

XX W0200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren P, Wang D;  
 Wang J, Wang Z, Wehrman-T, Xu C, Xue AJ, Yang Y, Zhang J;  
 Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB; AAI59228.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 such as central nervous system injuries -

XX Example 5; SEQ ID NO 3217; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 the encoded polypeptides (AAM38642-AAM42213) with nontropic,  
 immunosuppressant and cytostatic activity. The polynucleotides  
 in gene therapy. A composition containing a polypeptide or polynucleotide  
 of the invention may be used to treat diseases of the peripheral nervous  
 system, such as peripheral nervous injuries, peripheral neuropathy and  
 localised neuropathies and central nervous system diseases, such as

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 18, 2002, 08:14:06 ; Search time 24.71 seconds  
(without alignments)  
471.510 Million cell updates/sec

Title: US-09-777-921a-2

Perfect score: 2481  
Sequence: 1 MLRWLRDLPTRACDAEO.....VCISVYVYENKQILGVYQK 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents\_M.A.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1499	60.4	469	4	US-09-188-930-339
2	910	36.7	312	4	US-09-188-930-142
3	347	14.0	298	3	US-08-361-871-10
4	323.5	13.0	328	4	US-09-068-140A-15
5	311.5	12.6	289	4	US-09-068-140A-10
6	295	11.9	320	2	US-08-933-750C-12
7	295	11.9	320	4	US-08-933-750C-12
8	281	11.3	674	4	US-09-160-119-2
9	269.5	10.9	447	4	US-09-160-119-4
10	255	10.3	351	2	US-08-933-750C-19
11	255	10.3	351	4	US-08-933-750C-19
12	239.5	9.7	311	2	US-08-775-009-32
13	233.5	9.4	299	1	US-08-518-878B-56
14	233.5	9.4	309	2	US-08-518-878B-56
15	233.5	9.4	309	2	US-08-470-868A-51
16	233.5	9.4	309	2	US-08-470-868A-51
17	233.5	9.4	309	2	US-08-470-868A-51
18	233.5	9.4	309	3	US-09-210-681-51
19	233.5	9.4	309	3	US-08-946-719A-37
20	233.5	9.4	311	2	US-08-775-009-33
21	229	9.2	312	4	US-08-775-009-33
22	228	9.2	432	2	US-08-937-466-4
23	228	9.2	432	2	US-09-172-528-4
24	228	9.2	432	3	US-09-318-199-4
25	228	9.2	432	4	US-09-503-579-4
26	226	9.1	306	5	PCT-US94-09799-1
27	224	9.0	308	2	US-08-937-466-2

28	224	9.0	308	2	US-09-172-528-2	Sequence 2, Appli
29	224	9.0	308	3	US-09-318-199-2	Sequence 2, Appli
30	224	9.0	308	4	US-09-503-579-2	Sequence 2, Appli
31	198.5	8.0	303	1	US-08-518-878B-37	Sequence 37, Appli
32	198.5	8.0	303	1	US-08-294-522B-36	Sequence 36, Appli
33	198.5	8.0	303	2	US-08-807-861A-37	Sequence 37, Appli
34	198.5	8.0	303	2	US-08-470-868A-37	Sequence 37, Appli
35	198.5	8.0	303	3	US-09-210-681-37	Sequence 37, Appli
36	198.5	8.0	303	3	US-08-946-719A-37	Sequence 37, Appli
37	193.5	7.8	307	2	US-08-807-861A-56	Sequence 56, Appli
38	193.5	7.8	307	3	US-09-210-681-56	Sequence 56, Appli
39	193.5	7.8	307	3	US-08-905-223-320	Sequence 320, App
40	186	7.5	125	4	US-08-937-466-6	Sequence 6, Appli
41	179	7.2	256	2	US-09-172-528-6	Sequence 6, Appli
42	179	7.2	256	3	US-09-318-199-6	Sequence 6, Appli
43	179	7.2	256	4	US-09-503-579-6	Sequence 6, Appli
44	179	7.2	256	4	US-09-503-579-6	Sequence 6, Appli
45	178	7.2	149	1	US-08-100-874-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-188-930-339  
: Sequence 339, Application US/09188930A  
: Patent No. 6150502  
: GENERAL INFORMATION:  
: APPLICANT: Watson, James D.  
: APPLICANT: Strachan, Lorna  
: APPLICANT: Sleeman, Matthew  
: APPLICANT: Chrust, Rene  
: APPLICANT: Murison, James Greg  
: TITLE OF INVENTION: Compositions Isolated From Skin Cells  
: FILE REFERENCE: 11000.1011c1  
: CURRENT APPLICATION NUMBER: US/09/188,930A  
: CURRENT FILING DATE: 1998-11-09  
: NUMBER OF SEQ ID NOS: 348  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 339  
: LENGTH: 469  
: TYPE: PRT  
: ORGANISM: Mouse  
US-09-188-930-339

Query Match	60.4%	Score 1499;	DB 4;	Length 469;
Best Local Similarity	65.6%	Pred No 5.5e-138;		
Matches 274;	Conservative 73;	Mismatches 71;	Indels 0;	Gaps 0;
QY	60	EKIFTTGQVNGKGLDPEEFPMKYUKDHEKMKKLAFLKSLDKNNDKGKIEASEIVOSLQTLG	119	
Db	52	KOKIVAGGKDKLGGDGFEEFYHQLDHEKRLVFLKSLDKNNDKGKIEASEIVOSLQTLG	111	
QY	120	LTISEQAEILQSIDVDCGTTVDNNEWRDYFLNPVPTDTEETIRFWKHSTGTDIGDSLT	179	
Db	112	VKISEQAEILQSIDVDCGTTVDNNEWRDYFLNPVPTDTEETIRFWKHSTGTDIGDSLT	171	
QY	180	IPDETEDEKSKGOWRLLAGGTAGAVSRSTAPDLKIMQVHGSKSKDKNNFGGFR	239	
Db	172	VPDEFVEERQYGMWRHLVAGGAGAVSRCTAPDLKIMQVHGSKSKDKNNFGGFR	231	
QY	240	QMYEGGIRSLWRGNTNVIKAPETAVKFWAYDQYKLLTEECOKIGTGFPERFISGSMAG	299	
Db	232	QMYEGGIRSLWRGNTNVIKAPETAVKFWAYDQYKLLTEECOKIGTGFPERFISGSMAG	291	
QY	300	ATAQTFFYPMVKTRLAGVTGOYSGYDCAKRLKHEGKAFYKGYVNPILGIIPYAG	359	
Db	292	ATAQSSIFPMVKTRLAGVTGOYSGYDCAKRLKHEGKAFYKGYVNPILGIIPYAG	351	
QY	360	IDLAVTELLKSWLDNFADKSNVPGVWYVLCGCGALSTCCQLASYPPLAVTRTRMOAQL	419	

A. Haken to Weber et. al

Mon Aug 19 15:20:03 2002

us-09-777-921a-2.rpr

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 18, 2002, 08:23:51 ; Search time 43.26 Seconds  
(without alignments)  
1059.515 Million cell updates/sec

Title: US-09-777-921a-2

Perfect score: 2481  
Sequence: 1 MLRWLRDFAALPTAACQDAEQ.....VGISVYVYENKMTLGVTOK 477

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2373	95.6	475	2 T50686	peroxisomal Ca-dep
2	1103	44.5	588	2 T22688	hypothetical prote
3	991.5	40.0	531	2 G89667	protein F17E5.2 [1
4	988.5	39.8	587	2 T21074	hypothetical prote
5	852	34.3	479	2 T49871	peroxisomal Ca-dep
6	521	21.0	352	2 T01729	mitochondrial solu
7	520	21.0	332	2 T47703	Ca-dependent solu
8	506	20.4	358	2 T45934	hypothetical prote
9	500.5	20.2	500	2 T39385	probable mitochond
10	480.5	19.4	330	2 S26596	probable mitochond
11	475	19.1	348	2 D84798	probable mitochond
12	455.5	18.4	415	2 T48171	hypothetical prote
13	454.5	18.3	381	2 T51158	hypothetical prote
14	448	18.1	325	2 T04273	hypothetical prote
15	446	18.0	392	2 T05350	adenylate transloc
16	430.5	17.4	349	2 A40141	mitochondrial solu
17	429.5	17.3	294	2 T22145	hypothetical prote
18	428	17.3	326	2 B40141	mitochondrial solu
19	425.5	17.2	418	2 B96811	hypothetical prote
20	425.5	17.2	436	2 JQ4359	hypothetical prote
21	417.5	16.8	326	2 S57544	Btl protein precu
22	392	15.8	326	2 T37874	probable membrane
23	377.5	15.2	494	2 S57539	probable membrane
24	369	14.9	298	2 B43646	ADP,ATP carrier pr
25	368	14.8	298	1 S03894	ADP,ATP carrier pr
26	365	14.7	357	2 S46795	hypothetical prote
27	356	14.3	301	2 S21132	ADP,ATP carrier pr
28	355.5	14.3	386	2 S19174	ADP,ATP carrier pr
29	355.5	14.3	387	2 S14876	ADP,ATP carrier pr

ALIGNMENTS

RESULT 1  
150686  
peroxisomal Ca-dependent solute carrier [imported] - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50686  
R:Weber, F.E.; Minestrini, G.; Dyer, J.H.; Werder, M.; Boffelli, D.; Compassi, S.; We  
Proc. Natl. Acad. Sci. U.S.A. 94, 8509-8514, 1997  
A:Title: Molecular cloning of a peroxisomal Ca2+-dependent member of the mitochondria  
A:Reference number: 225180; MUID:97385133  
A:Accession: T50686  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-475 <WEB>  
A:Cross-references: EMBL:AF004161; PIDN:AAB69156.1  
C:Keywords: peroxisome

Query Match 95.6%; Score 2373; DB 2; Length 475;  
Best Local Similarity 95.2%; Pred. No. 4.8e-149;  
Matches 454; Conservative 12; Mismatches 9; Indels 2; Gaps 1;

Qy	1	MLRWLRDFAALPTAACQDAEQPRYETFLFQALDRNGDVVDIGELQELGRNIGILGQDAE	60
Db	1	MLRWLRGFLVPTAACGAEPPRYETFLFQALDRNGDVVDIRELQELGKLSIGILGQDAE	60
Qy	61	EXITFTGDNKDGKLDFFEFHMYLKDHEKMKLAFKSLDKNNDKGKIEASEIVQSLQTLGL	120
Db	61	EXITFTGDNKDGKLDFFEFHMYLKDHEKMKLAFKSLDKNNDKGKIEASEIVQSLQTLGL	120
Qy	121	TISEQQAELILQSIDVGDGTMVTVDNNEWRDYLFPNPVTDEIIRFWKHSKGIDGSLTI	180
Db	121	TISEQQAELILQSIDAGDGTMTVDNNEWRDYLFPNPVADIEIIRFWKHSKGIDGSLTI	180
Qy	181	POEFTDEKSKQWROLLAGGIAGAVSRTSTAPLDRLKIMQVHGSKDKMNTFGFRQ	240
Db	181	POEFTDEKSKQWROLLAGGIAGAVSRTSTAPLDRLKIMQVHGSKDKMNTFGFRQ	240
Qy	241	MYKEGIRSLMRGNTNVIKIPATAVKFWAYEQYKLLTTEGOKICTFERFISGSNAGA	300
Db	239	MIKEGVRSLMRGNTNVIKIPATAVKFWAYEQYKLLTTEGOKICTFERFISGSNAGA	298
Qy	301	TAQTIYPMVKNKTRVLAVKGTGOYSGIYDCAKKILKHEGLGAFYKGVYVPLLGIIPYAGI	360
Db	299	TAQTIYPMVKNKTRVLAVKGTGOYSGIYDCAKKILKHEGLGAFYKGVYVPLLGIIPYAGI	358
Qy	361	DLVAYELLKSKWLDNFAKDSVNPVYVLLGCGALSSTCGQLASYPALVTRMQAAMLE	420
Db	359	DLVAYELLKSKWLDNFAKDSVNPVYVLLGCGALSSTCGQLASYPALVTRMQAAMLE	418
Qy	421	GSPQLNMYGLFRRIISKEGICPLVGRITPMPKVLPAVGISVYVYENKMTLGVTOK	477
		1:	

Db 419 CAPOLNMGVLFRRIRISREGLPLRGITPDMFNKVLPAVGISVYVYENMKOTILGVTK 475

RESULT 2

T22688

hypothetical protein F55A11.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T22688

R:Kershaw, J.

submitted to the EMBL Data Library, May 1996

A:Reference number: 219600

A:Accession: T22688

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-588 <N1>

A:Cross-references: EMBL:272511; PIDN:CAA96658.1; GSPDB:GN00023; CESP:F55A11.4

C:Genetics:

A:Gene: CESP:F55A11.4

A:Map position: 5

A:Introns: 24/1; 77/2; 277/3; 434/3; 474/1; 511/3; 531/3

Query Match 44.5%; Score 1103; DB 2; Length 588;

Best Local Similarity 45.7%; Pred. No. 3.9e-65;

Matches 210; Conservative 104; Mismatches 142; Indels 4; Gaps 3;

Qy 19 EOPTRYETLFOALDRNGDGVVDIGELGRNKGIPGLQDAEEKITFTGVDNKGKLDPE 78

Db 73 EKERQIRDIYDRLDIDNDGTDIDRLTALKHETPHANLAPVIMSKSPDDEGRVDFY 132

Qy 79 EFMKYLKDKHEKMKLAFKSLDKNNDGKIEASEIVOSLOTGLTISEQQAELILQSIDVG 138

Db 133 SFSSVYLENEOKLAEMFADNRHNDGLVDVENKYNCKDYGVLDDHKAQHVNMKMDQTG 192

Qy 139 TMTVDNEMRDYFLFNPVTDIEIIRFWKSTGIDIGDLSLTIPDETEDEKSKGOWROL 198

Db 193 SASVDLEFEPFPMPLPSSDLKIDVDFWRNLIIIDGESQIPEDFSQEQBQGIWRRHL 252

Qy 199 LAGIAGAVSTRTAPLDRUKIMQVHSGSKDKMNIFFGFRQVMEGGIRSLWRNGSTNV 258

Db 253 VAGAGAVSTRTAPLDRUKIMQVHSGSKDKMNIFFGFRQVMEGGIRSLWRNGSTNV 312

Qy 259 IKIAPETAVFWAYEYKLLTTEE--GQKIGTFERFISGSMAGATAQTFIYPMEVMKTRL 316

Db 313 IKIAPESAIKFCYDQLKRLIQKKNEISTFERLCAGSAAGAISQSTIYPMEVMKTRL 372

Qy 317 AVGTGQYS-GIYDCAKKILKHEGAFYGVNPLIGIIPYAGIDLAVYELLKSYWLDN 375

Db 373 ALRKTGQDRGIHFHAKMTKEGICFCYGLPDLNIGIIPYAGIDLAIYETLKRITYRY 432

Qy 376 FAKDSVNPVGYVLLGCCALSSCGQLASYPALVTRMQAAMLEGSPQLN-MVGLFERRI 434

Db 433 YETNSPFGVLLACGTCSSCGQLSSPFPALVTRMQAAMLEGSPQLN-MVGLFERRI 492

Qy 435 ISKEGIPGLYRGITPDMFNKVLPAVGISVYVYENMKOTILGV 474

Db 493 LQNSGVYGFYRGITPDMFNKVLPAVGISVYVYENMKOTILGV 532

RESULT 3

G89667

protein F17E5.2 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001

C:Accession: G89667

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; WUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: G89667

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-531 <STO>

A:Cross-references: GB:chr\_X; PIDN:CAA90761.1; PID:q3876012; GSPDB:GN00028; CESP:F17E5.2

C:Genetics:

A:Gene: F17E5.2

A:Map position: X

Query Match 40.0%; Score 991.5; DB 2; Length 531;

Best Local Similarity 42.3%; Pred. No. 7.7e-58;

Matches 203; Conservative 92; Mismatches 138; Indels 47; Gaps 8;

Qy 16 QDAEOPTRYETLFOALDRNGDGVVDIGELGRNKGIPGLQDAEEKITFTGVDNKGKLD 75

Db 11 QNISRLSIRDMYDRLDADNDGSDIRDLQAL-SLOAHIPASVAPKLLERKSEHSORV 69

Qy 76 DFEFPMKYLKDKHEKMKLAFKSLDKNNDGKIEASEIVOSLOTGLTISEQQAELILQSID 135

Db 70 TYADFTNYVIAHEARLAEVFDKIDLSNGVDMAEIKSKYCKENGVLDDQKAMSVKMD 129

Qy 136 VDGTMVDNEMRDYFLFNPVTDIEIIRFWKSTG-----IDIGDSL 178

Db 130 QSGSSSVNLNEFQDFMLLYPSTDRHMDVDFWRNLVCTCLSESNRFTQNFOLIIDIGDG 189

Qy 179 TIPDETEDEKSKGOWROLLAGGIAGAVSTRTAPLDRUKI-----M 221

Db 190 QVPEDFTQELLSGVWRRHLVAGVAGASRTCTAPDRIKYLVQVYLVHLLFHFMLKA 249

Qy 222 MOVHSGSKDKMNIFFGFRQVMEGGIRSLWRNGSTNVIKIAPETAVFWAYEYKLLTTEE 281

Db 250 LQVNSTKTKNGVYSCVHLHAEGGKISFWRGNGINVIKIAPEASAKMFCYDQIKRMQOE 309

Qy 282 --EGOKIGTFERFISGSMAGATAQTFIYPMEVMKTRLAVGKTCQY-SGIYDCAKKILKHE 338

Db 310 YKGAELSTIERLAGSSAGAISQSTIYPMEVMKTRLALARTQLOKGMFHFHAKMYTKE 369

Qy 339 GLGAFYKGVNPLIGIIPYAGIDLAVYELLKSYWLDNFAKDSVNPVGYVLLGCCALSSTC 398

Db 370 GKICFYKTLNPLIGIIPYAGIDLVYESLSMT-TKIYITEHPGVLLACGTCSSTC 428

Qy 399 GOLASYPALVTRMQAAMLEGSPQLN-----MVGLFRRIRISREGLPLRGITPDMFNK 453

Db 429 GOLASYPALVTRMQAAMLEGSPQLN-----SPKNSSTOPDMVGFQKHILQTEGFTGLYRGITPDMFNK 485

RESULT 4

T21074

hypothetical protein F17E5.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T21074

R:McMurray, A.

submitted to the EMBL Data Library, August 1995

A:Reference number: 219368

A:Accession: T21074

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-587 <N1>

A:Cross-references: EMBL:250873; PIDN:CAA90761.2; GSPDB:GN00028; CESP:F17E5.2

C:Genetics:

A:Gene: CESP:F17E5.2

A:Map position: X

A:Introns: 21/1; 48/3; 74/2; 111/3; 183/2; 384/1; 432/3; 503/1; 541/3

Query Match 39.8%; Score 988.5; DB 2; Length 587;

Best Local Similarity 42.3%; Pred. No. 1.4e-57;

Matches 202; Conservative 91; Mismatches 133; Indels 47; Gaps 8;

Qy 19 EOPTRYETLFOALDRNGDGVVDIGELGRNKGIPGLQDAEEKITFTGVDNKGKLDPE 78